

Supplementary Material

The Personal Genome Browser: Visualizing functions of genetic variants

Liran Juan, Mingxiang Teng, Tianyi Zang, Yafeng Hao, Zhenxing Wang, Chengwu Yan, Yongzhuang Liu, Jie Li, Tianjiao Zhang, and Yadong Wang*

Center for Bioinformatics, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China

*Corresponding Author: Yadong Wang, ydwang@hit.edu.cn

Figure S1	Combination of individual genome panel and reference panel
Figure S2	Symbols and shapes for individual genome visualization
Figure S3	Input personal genome variants files into PGB
Figure S4	Basic operations of the PGB
Figure S5	The PGB track management
Figure S6	Dense and pack modes of the five track types in the PGB
Figure S7	Example uses
Figure S8	The correlations and time costs of accurate reads number method and approximate file offset method.
Table S1	Integrated databases
Table S2	Variant functional significance evaluated by PGB, SIFT and PolyPhen2 scores
Table S3	Variant effects
Table S4	Integrated phenotype databases
Table S5	Comparison of GFF3 and GRF/GDF formats
Visual elements	Visual elements of the five types of tracks
Examples	Examples
Data privacy	Data privacy and security
User guide of the PGB	User guide of the PGB



Figure S1. Combination of individual genome panel and reference panel. (A) Clicking on the ‘Merge’ button on the individual genome panel controls, the GMP tracks are displayed in the reference genome panel, and a group of controls appear on the right side. (B) Clicking on the ‘<=>’ button, the GMP tracks are removed from the reference genome panel and the individual genome panel is restored. (C) Clicking on the ‘Gear’ button, users can set up the GMP tracks, which is the same as the ‘setting’ button in individual genome panel controls. (D) Users can use the ‘X’ button to remove GMP tracks, which is the same as the ‘close’ button in individual genome panel controls.

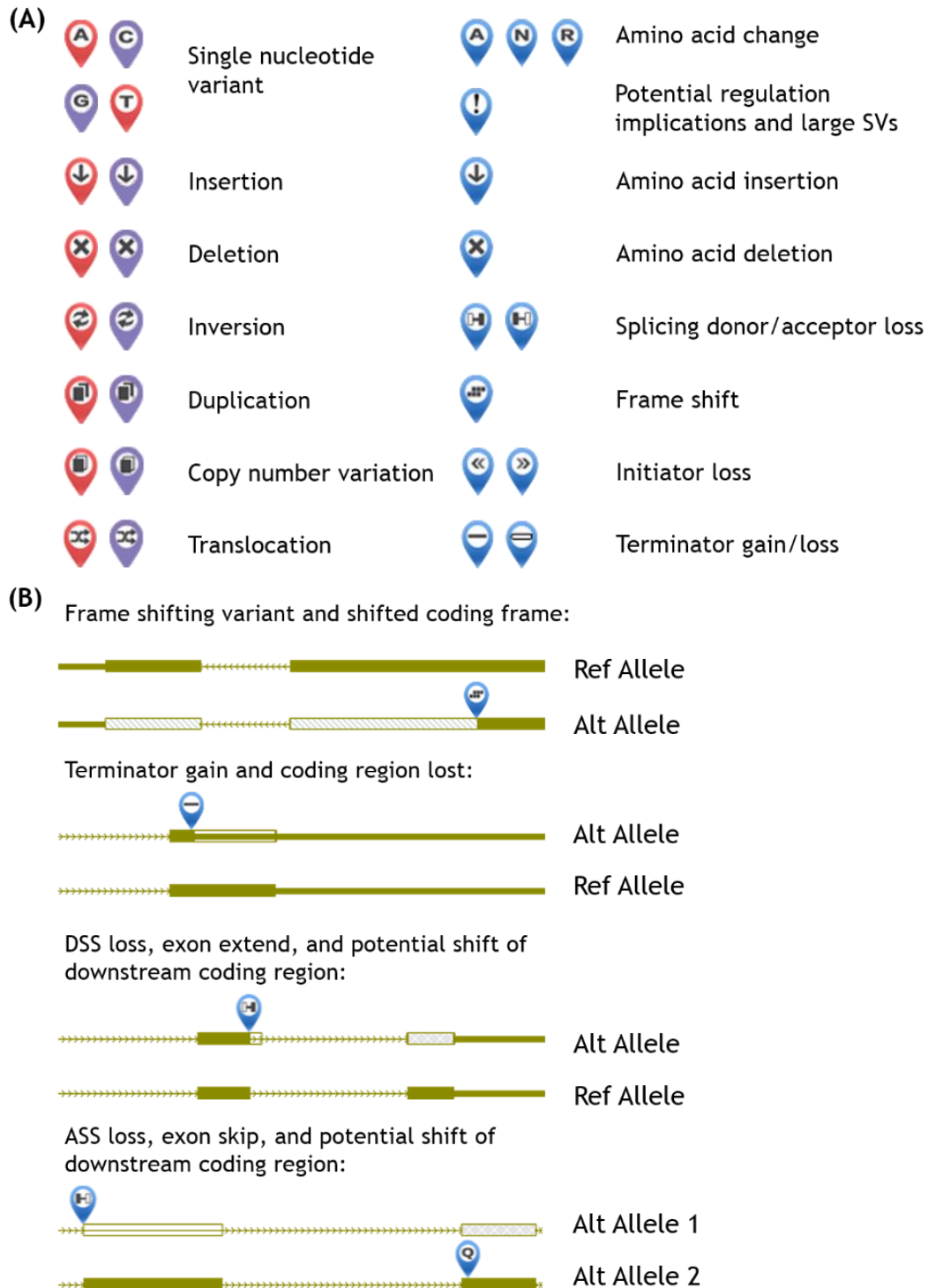


Figure S2. Symbols and shapes for individual genome visualization. (A) Symbols for the individual genome, with red/purple ones representing genetic variants and blue ones representing amino acid changes. (B) Shapes extended from the UCSC genome browser (1) characterizing individual functional effects on the molecular level.

The screenshot displays the PGB interface with the following components and annotations:

- Upload Individual Genome:**
 - Track Name:
 - File Format:
 - (A)** Radio buttons for "URL:" (selected) and "Data File:".
 - URL: (Recommended)
 - Data File: No file selected. (<20MB)
 - Index File: No file selected.
 - (B)** "Submit" and "Example" buttons.
- Note:**
 - The track name cannot be same to existing tracks and cannot start with '_'.
 - The data must be compressed and indexed by [Tabix](#).
 - User can either provide accessible data URL link (recommended) or locally upload data and index files.
 - For URL link, the data file and index file must be placed together, the index file name should be data file name + '.tbi'.
 - The data file should follow standard format. Check the format standard of [vcf](#) and [gvf](#).
 - The header line '#CHROM...' is required for VCF file.
- Personal Genomes:**
 - Buttons: Venter, Watson, **test_1** (orange), **(D)** (delete 'x')
 - (C)** indicates built-in and user-uploaded genomes.
- Built-in Genomes from 1000 Genome Project:**
 - Buttons: ASW, CEU, CHB, CHS, CLM, FIN, GBR, IBS, JPT, LWK, MXL, PUR, TSI, YRI.
 - (E)** indicates built-in genomes.
- Select Individual:**
 - Radio buttons: NA11992, NA11993, NA11994, **(G)** NA11995 (selected).
 - (F)** indicates the track name radio button.

Figure S3. Input personal genome variants files into the PGB. **(A)** Upload users' individual genome variants data by either providing data file URLs or uploading the local data file. **(B)** Click the 'Example' button to fill an example URL into blanks. **(C)** Built-in and users' uploaded personal genomes, blue labels represent built-in genomes, and orange ones are users' uploaded genomes. **(D)** Click the 'x' button to delete users' uploaded (orange) personal genomes. **(E)** Built-in genomes from the 1000 Genomes Project. **(F)** Click the track name radio button to show the selected individual in **(G)**. **(G)** Select an individual genome in the file. The selected individual genome will be loaded into the individual genome panel.

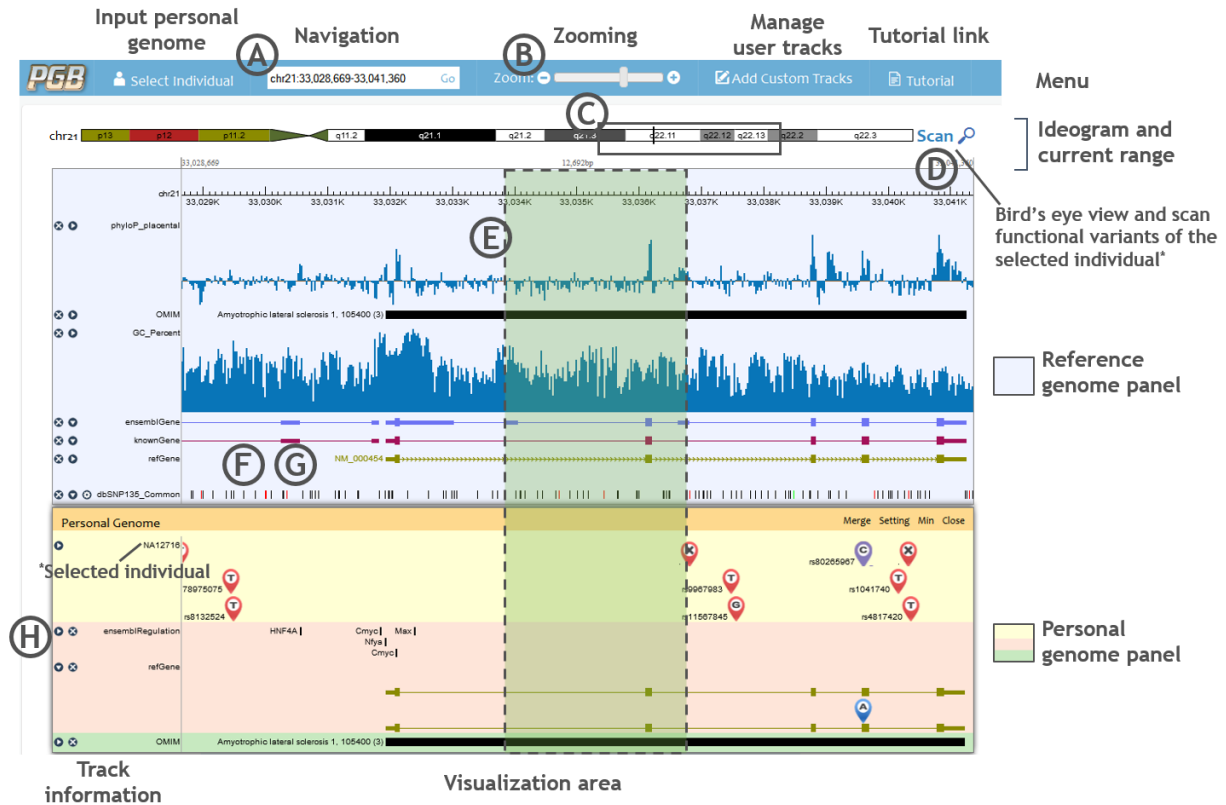


Figure S4. Basic operations of the PGB. **(A)** Navigation by specifying genomic coordinates or gene symbols in the input textbox. **(B)** Zooming by clicking ‘+’ and ‘-’ buttons or dragging the slider. **(C)** Navigating to the interested region by clicking and dragging the cursor on the chromosome ideogram. **(D)** Using the whole genome bird’s eye view to select and navigate to a region. **(E)** Holding the 'shift' and using the cursor to select a region to zoom in. **(F)** Holding the 'z' key and scrolling the mouse wheel to zoom out/in. **(G)** Dragging the visualization area in either direction to move left or right within the current zoom level. **(H)** Clicking the ‘x’ button to close a track, and clicking the triangle button to change the track display mode.

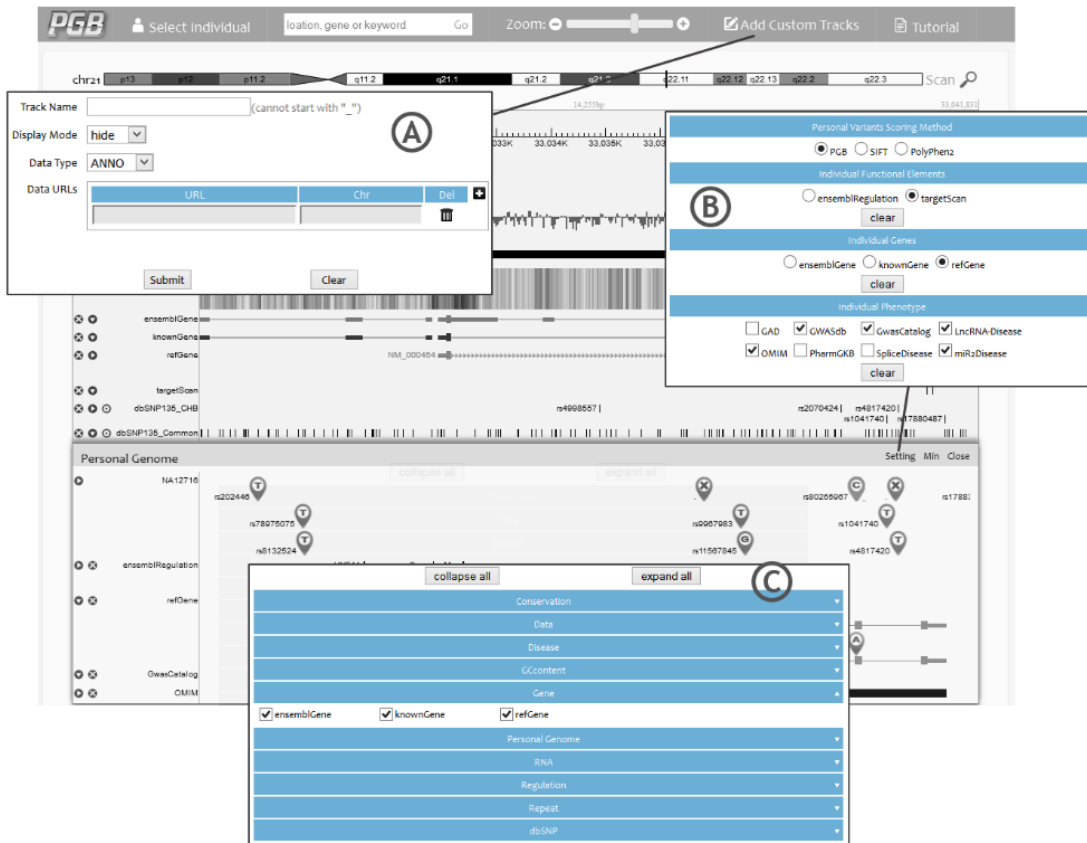


Figure S5. The PGB track management. **(A)** Adding custom tracks to the reference genome panel. **(B)** Managing tracks in the individual genome panel. **(C)** Managing tracks in the reference genome panel.

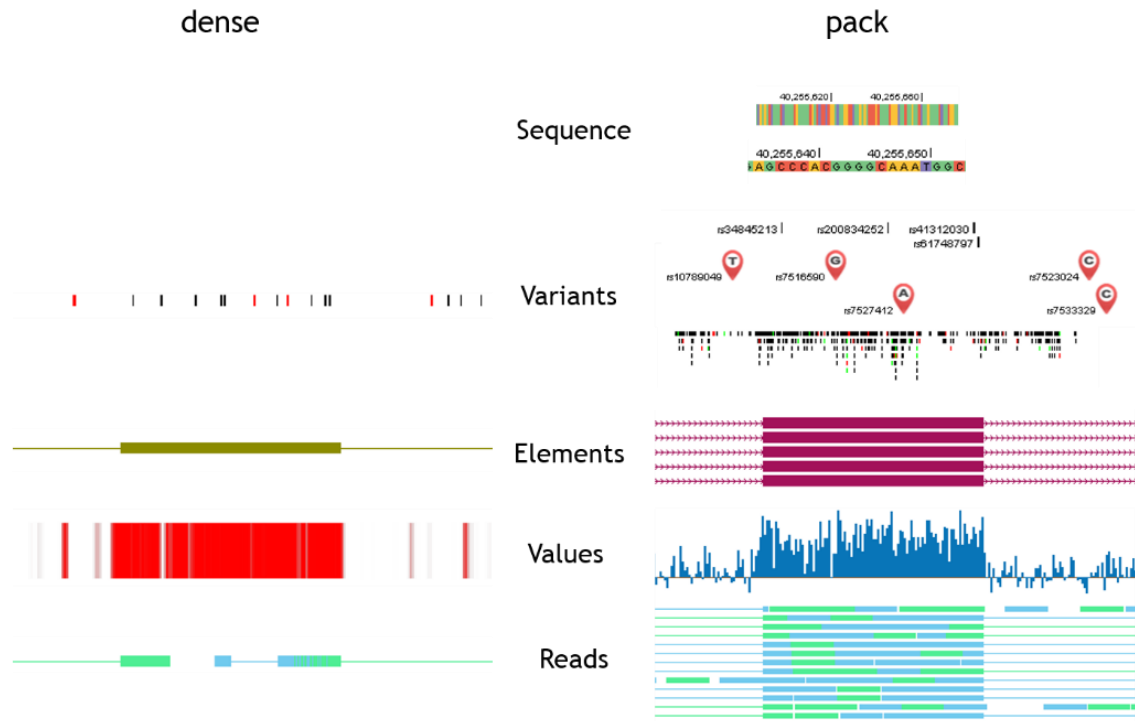


Figure S6. Dense and pack modes of the five track types in the PGB. Details about the visual elements of the five tracks types are described in the “Visual elements” section.

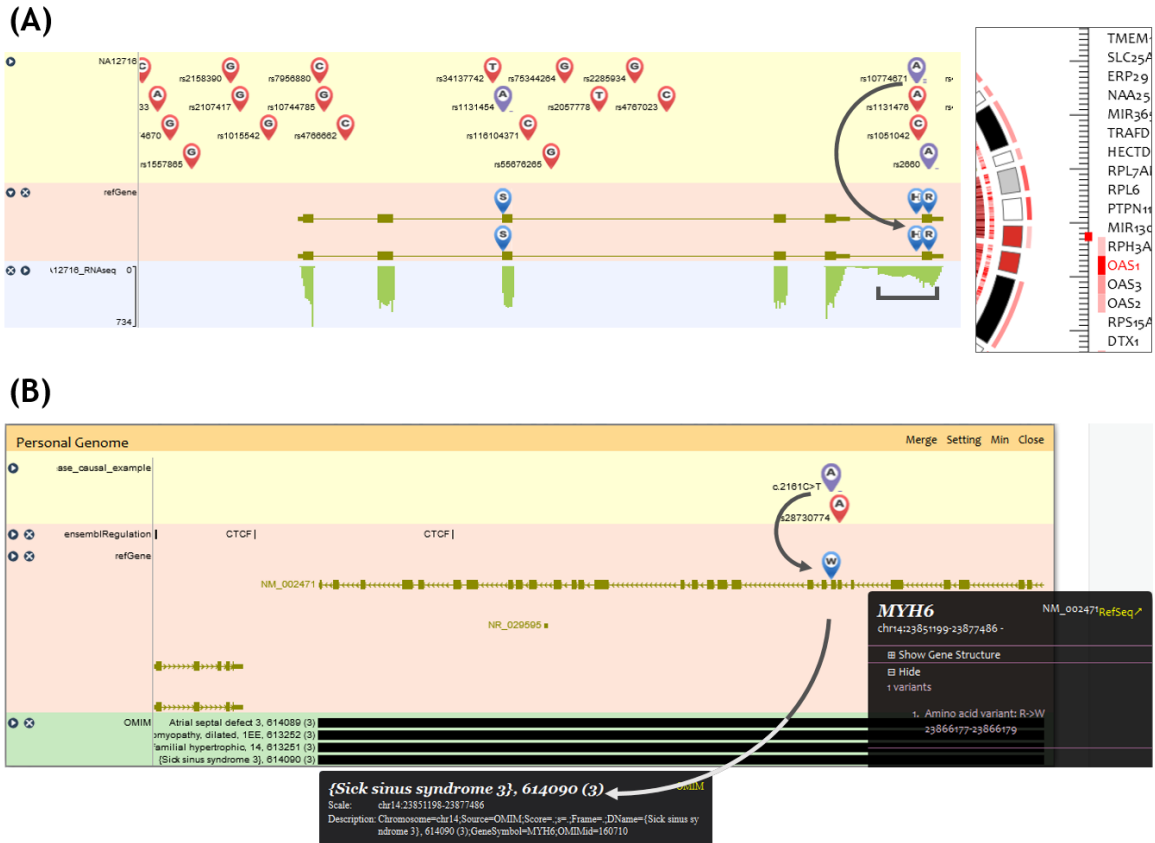


Figure S7. Two Examples. (A) A splicing acceptor loss is caused by the variant rs10774671 through the inference. Irregular expression of the exon is observed from the RNA-seq data of the same individual, which validates the function of the variant. **(B)** A disease case study in the PGB. A rare variant in the MYH6 associated to the Sinus sick syndrome (SSS) is visualized as non-synonymous, which is consistent with the original study. The details of the two examples are described in the “Examples” section.

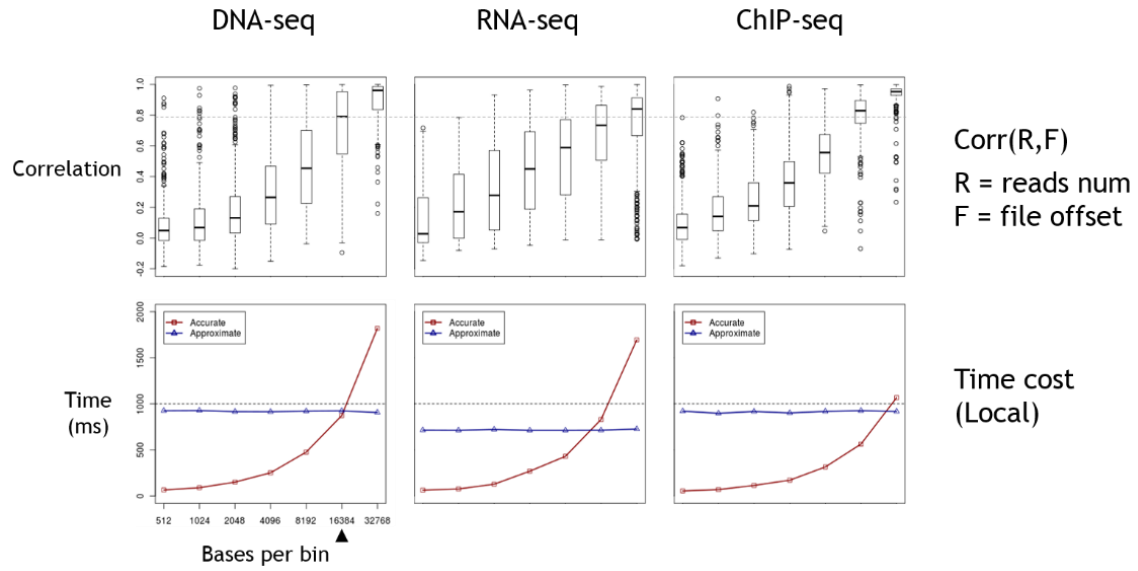


Figure S8. The correlations and time costs of accurate reads number and approximate file offset methods.

Table S1. Integrated databases

Built-in track	Data format	Track type	Group	Source
Hg19	Fasta	Sequence	Basic	UCSC
refseqGene ^a	ANNO	Elements	Genes	UCSC
knownGene ^a	ANNO	Elements	Genes	UCSC
ensemblGene ^a	ANNO	Elements	Genes	UCSC
Cytoband	cytoBand	--	Basic	UCSC
HGNC Genes	Txt	--	Basic	HGNC
dbSNP138	VCF	Variants	dbSNP	NCBI
1000 genome phase1	VCF	Variants	Individual genomes	1000 Genomes Project
Watson genome	GVF	Variants	Individual genomes	Ensembl
Venter genome	GVF	Variants	Individual genomes	Ensembl
phyloP	BigWig	Values	Conservation	UCSC
phastCons	BigWig	Values	Conservation	UCSC
GC content	BigWig	Values	Conservation	UCSC
repeatMask	Bed	Elements	Repeat	UCSC
Human RNA	Bed	Elements	RNA	UCSC
ensemblRegulation ^b	GRF	Elements	Regulation	Ensembl
targetScan ^b	GRF	Elements	Regulation	targetScan
NA12716_RNAseq	BAM	Reads	Data	1000 Genomes Project
Phenotypic databases ^c	GDF	Elements	Disease	Multiple databases
Outgroup genomes	FASTA	Sequence	Pairwise alignments	UCSC

^a Gene annotations tracks, which can be annotated and visualized in individual genome panel. These tracks are in the 'ANNO' format. It is the same as the BED format, except the third column recording the Gene symbol.

^b Functional element annotations tracks, which can be annotated and visualized in individual genome panel. These tracks are in the 'GRF' format, a customized version of the GFF3. Detail of the GRF can be found in Table S5.

^c Tracks include eight disease and drug databases. These databases record the relationships between disease/drug and genomic elements. Table S4 lists the detail of these databases. Since the database system is not deployed in the PGB, we also customized a special version of the General Feature Format for these disease/drug features, with a slightly variation in column 3 and 9. Table S5 compares the protocol of the GFF3 (2) and its extension: the General Disease/Drug Format (GDF).

a, b, and c tracks accept remote or local additional customized annotations in given formats.

Table S2. Variant functional significance for the PGB, the SIFT and the PolyPhen2 prediction scores

Functional significance	PGB	SIFT	PolyPhen2
Level 1	AA substitution, INS, DEL	$-\log_{10}(\text{SIFT score}) > 1.3^*$	$-\log_{10}(1 - \text{PolyPhen2 score}) > 1.3$
Level 2	Stop gain/loss	$-\log_{10}(\text{SIFT score}) > 2.3$	$-\log_{10}(1 - \text{PolyPhen2 score}) > 2.3$
Level 3	ASS loss, DSS loss	$-\log_{10}(\text{SIFT score}) > 3.3$	$-\log_{10}(1 - \text{PolyPhen2 score}) > 3.3$
Level 4	Frame shift, start codon loss	$-\log_{10}(\text{SIFT score}) > 4.3$	$-\log_{10}(1 - \text{PolyPhen2 score}) > 4.3$
Level 5	SV in coding region	$-\log_{10}(\text{SIFT score}) > 5.3$	$-\log_{10}(1 - \text{PolyPhen2 score}) > 5.3$

* $-\log_{10}(0.05) \approx 1.30103$

Table S3. Variant effects

Functional role of variant	Variant effect displayed in gene/functional elements tracks of individual genome panel (Figure S2)	Variant type
Regulatory variation	Exclamation mark (and using red colour to highlight the functional elements containing variant)	All
Synonymous SNV	No amino acid changing displayed	SNV
Non-synonymous SNV	Amino acid substitution	SNV
Stop Gain	3' Coding region lost	SNV/INDEL
Stop Loss	5' Coding region extend	SNV/INDEL
Start Codon Loss	Initiation postpone	SNV/INDEL
Frame shift	Coding frame shift	$\sum \text{INDEL} \% 3 \neq 0$ *
Frame shift	Coding frame recovery	$\sum \text{INDEL} \% 3 = 0$ *
Insertion	Amino acid insertion	$\text{INS} \% 3 = 0$ *
Deletion	Amino acid deletion	$\text{DEL} \% 3 = 0$ *
ASS Loss	Exon skip	SNV/INDEL
DSS Loss	Exon extend	SNV/INDEL
SV in coding region	Exclamation mark	Structural variation

* % refers to the MOD function

Table S4. Integrated phenotypic databases

Database name	Data source	Record	Reference
OMIM	Publications	4837	(3)
GAD	Integrated	112538	(4)
lncRNA-Disease	Publications	513	(5)
miR2Disease	Publications	2925	(6)
GWAS Catalog	Publications	9421	(7)
GWASdb	Integrated	208502	(8)
PharmGKB	Publications	17960	(9)
SpliceDisease	Publications	3415	(10)

Table S5. Comparisons of the GFF3 and GRF/GDF formats

Column #	GFF3	GDF	GRF
1	Seqid	Chromosome	Chromosome
2	Source	Source	Source
3	Type	Type	Type
4	Start	Start	Start
5	End	End	End
6	Score	Score	NA
7	Strand	NA	Strand
8	Phase	NA	NA
9	Attributes	Attributes	Attributes
9-1	NA	DName	FactorName
9-2	NA	GeneSymbol	GeneSymbol
9-3	NA	dbSNPid	Motif
9-4	NA	Original Access ID	Reference
9-5	NA	Alt names	NA
9-6	NA	DOID	NA
9-7	NA	Reference	NA

Visual elements

Genomic data in different formats are displayed by five track types in the PGB, including sequences, variants, elements, values and reads. The sequence type has one display mode. Other four types have two display modes: the dense mode and the pack mode (Figure S6).

In the sequence track, the reference genome is a sequence of bases, which are denoted by different colors and letters.

In variants tracks, the genomic location and the alleles are fundamental information about genetic variants. In the individual genome panel, symbols with different colors and labels are used to indicate the location and alleles of a variant. Symbols in red indicate non-functional variants (predicted), and purple ones indicate possible functional variants. Symbols of large scale SVs are placed in the middle of the SV region. Similar with the squish mode of the UCSC Genome Browser (1), when there are a large number of variants within the browsing region, the symbols are replaced by short vertical lines (or bands if the variants cover large regions). Different types of variants are denoted by different colors: SNVs are black, insertions are red, deletions are green, and SVs are blue. Users can then get an overall view of the browsing region. Variants tracks in dense mode are also displayed in this color scheme.

Elements tracks describe the genomic features in discrete regions. The attributes of the features, such as exon, intron, coding region, direction, etc., are denoted by different shapes.

Values tracks describe features of genomic positions/regions, including conservation, gene expression level, reads enrichment, DNA methylation, histone modification, and GC content. The heatmap and histogram are adopted to display the values of these features.

In reads tracks, high-throughput sequencing reads data are displayed in two different views. If each read is legible, all reads are shown separately in the browser. If reads are illegible, reads coverage are shown using the heatmap or the histogram.

Examples

In this section, we show two examples to demonstrate our personal genome browser.

We employ the NA12716, an individual from the CEU population of the 1000 Genomes Project to show the PGB's functionality. The functional inspection of this individual genome suggests that the *OAS1* gene on chr12q24.13 is a gene with high risks. As shown in Figure S7A, an SNV *rs10774671* is located at the splicing acceptor of the last exon of *OAS1* (Transcript ID: NM_016816). The PGB not only marks up the location and functional category of this variant, but also indicates it leads to a splicing interruption, i.e., the functional role this variant plays. This discovery is consistent with the high-throughput RNA sequencing experiment dataset from the Geuvadis project (<http://www.geuvadis.org>).

The other example is a disease case study involving multiple variants (Figure S7B). In this study, two variants of *rs28730774* and *c.2161C>T* in *MYH6* gene were found to be associated with high risks of *sick sinus syndrome* (SSS) (11). The stronger association observed was with the missense mutation *c.2161C>T*. Based on the knowledge in the OMIM database, the PGB displays the association between SSS and the two variants correctly. In the PGB, SNV *rs28730774* and mutation *c.2161C>T* are shown in the variants track by distinct variant symbols, a conceptual amino acid substitution *p.Arg721Trp* caused by mutation *c.2161C>T* is displayed in the gene tracks. The PGB can intuitively display the functional consequences of the two variants. The variant *c.2161C>T* changing the amino acid of *MYH6* product is illustrated and *rs28730774* does not. This confirms conclusions by Holm et al (11).

Data privacy and security

The data privacy and ownership are protected in the PGB through following approaches:

- a) A firewall has been deployed in the PGB server to protect the server from attack.
- b) The PGB uses session mechanism to protect the data privacy and ownership. The interaction between each user and the PGB server is based on a unique session. Within a session, user specific information cannot be accessed by the other users. And all user specific information will be removed when the session expires.
- c) The linked user data are only temporally loaded in memory rather than in server disk. The user uploaded files are stored in temporary folders in the server, and can be deleted by the user via PGB. When the user session expires, all the user data are deleted.
- d) The PGB uses the HTTPS protocol to provide encryption during data transmission. Users can visit <https://www.pgbrowser.org:8443/> to access PGB.
- e) In addition, we have released the PGB as a separate package for local installation. The package is available at http://www.pgbrowser.org/pgb_1.0_local_installation.tar.gz. Users can use the local PGB to visualize private/sensitive data, such as patient data and lab specific data.

References

1. Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M. and Haussler, D. (2002) The human genome browser at UCSC. *Genome Res*, **12**, 996-1006.
2. Eilbeck, K., Lewis, S.E., Mungall, C.J., Yandell, M., Stein, L., Durbin, R. and Ashburner, M. (2005) The Sequence Ontology: a tool for the unification of genome annotations. *Genome Biol*, **6**, R44.
3. Hamosh, A., Scott, A.F., Amberger, J.S., Bocchini, C.A. and McKusick, V.A. (2005) Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res*, **33**, D514-517.
4. Becker, K.G., Barnes, K.C., Bright, T.J. and Wang, S.A. (2004) The genetic association database. *Nat Genet*, **36**, 431-432.
5. Chen, G., Wang, Z., Wang, D., Qiu, C., Liu, M., Chen, X., Zhang, Q., Yan, G. and Cui, Q. (2013) LncRNADisease: a database for long-non-coding RNA-associated diseases. *Nucleic Acids Res*, **41**, D983-986.
6. Jiang, Q., Wang, Y., Hao, Y., Juan, L., Teng, M., Zhang, X., Li, M., Wang, G. and Liu, Y. (2009) miR2Disease: a manually curated database for microRNA deregulation in human disease. *Nucleic Acids Res*, **37**, D98-104.
7. Hindorff, L.A., Sethupathy, P., Junkins, H.A., Ramos, E.M., Mehta, J.P., Collins, F.S. and Manolio, T.A. (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A*, **106**, 9362-9367.
8. Li, M.J., Wang, P., Liu, X., Lim, E.L., Wang, Z., Yeager, M., Wong, M.P., Sham, P.C., Chaoock, S.J. and Wang, J. (2012) GWASdb: a database for human genetic variants identified by genome-wide association studies. *Nucleic Acids Res*, **40**, D1047-1054.
9. Whirl-Carrillo, M., McDonagh, E.M., Hebert, J.M., Gong, L., Sangkuhl, K., Thorn, C.F., Altman, R.B. and Klein, T.E. (2012) Pharmacogenomics knowledge for personalized medicine. *Clinical pharmacology and therapeutics*, **92**, 414-417.
10. Wang, J., Zhang, J., Li, K., Zhao, W. and Cui, Q. (2012) SpliceDisease database: linking RNA splicing and disease. *Nucleic Acids Res*, **40**, D1055-1059.
11. Holm, H., Gudbjartsson, D.F., Sulem, P., Masson, G., Helgadóttir, H.T., Zanon, C., Magnusson, O.T., Helgason, A., Saemundsdóttir, J., Gylfason, A. *et al.* (2011) A rare variant in MYH6 is associated with high risk of sick sinus syndrome. *Nat Genet*, **43**, 316-320.

User guide of the PGB

Accessibility

The PGB is recommended to be accessed by Google Chrome, and works smoothly as well with other web-browsers, including Mozilla Firefox, Safari, Microsoft Internet Explorer (Version 10 or later), Opera, etc. HTML5 Canvas is used as the graphic engine to plot the visual elements, thus some earlier versions of web-browsers that do not support HTML5, such as IE6-9, are not compatible with the PGB.

The PGB is available at <http://www.pgbrowser.org>.

The latest PGB tutorial is available at <http://www.pgbrowser.org/tutorial.html>.

The PGB local installation package is available at http://www.pgbrowser.org/pgb_1.0_local_installation.tar.gz.

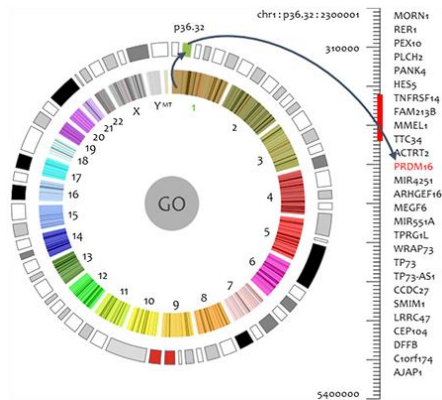
Basic operations of the PGB

1. Navigating the browser region by the input textbox (Figure S4A)
 - Navigating to the interested region by specifying chromosome coordinates in the input textbox:

Input chromosome coordinates	Valid or not
chr1:1234-5678	YES
chr1:1,234-5,678	YES
Chr1:1234-5678	NO
chr1 1234 5678	NO
1:1234-5678	NO

- Navigating to the gene corresponding region by inputting the prefix of gene symbols in the input textbox.

The PGB can list top-5 candidate genes that start with the input prefix in lexicographical order. Users can navigate to the corresponding region by clicking on the listed gene symbols. The acceptable input gene symbol is case-independent.
2. Zooming out or zooming in the browsing region (Figure S4B)
 - Clicking the '-' or '+' buttons.
 - Moving the slider left or right.
 3. Navigating to the selected browsing region by clicking and dragging the cursor on the chromosome ideogram (Figure S4C).
 4. Navigating to the selected browsing region by the whole genome bird's eye view (Figure S4D)



Move the mouse to the visualization area, when the mouse icon changes to ‘grabber’, then,

5. Holding the 'shift' and using the cursor to select a region to zoom in (Figure S4E).*
6. Holding the 'z' key and scrolling the mouse wheel to zoom out/in (Figure S4F).*

*These operations may be invalid if the hotkey is occupied by browser plugins or other programs.

7. Dragging the visualization area in either direction to move left or right within the current zoom level (Figure S4G).
8. Clicking the ‘x’ button to hide a track. Clicking the ‘right/down triangle button’ to change the track display mode (dense/pack) (Figure S4H).

Data input

Users can click the Select Individual button on the menu bar to submit input data to the PGB (Figure S3). The input of the PGB is an individual genome variants file. The PGB accepts Bgzip/Tabix compressed/indexed VCF and GVF files. Users can either provide accessible data URLs or upload local data and index files. Using URL link is recommended. The PGB guarantees input data can only be accessed and viewed by its owner. The PGB will NOT record any information of users’ data. The uploaded files will be automatically deleted as soon as the session expired.

Users can also browse built-in genomes for comparisons. We currently hold over a thousand individual genomes in the PGB server. Most of them are from the 1000 genomes project.

Variants track parameter setting

Currently, variants tracks in the reference genome panel allow users to set parameters by the gear button. When clicking on the gear button, the track parameters setting window pops up for setting three parameters of the quality threshold, filters, and samples.

- ‘Quality Threshold’ specifies a value used to choose variants whose quality is not less than the given value.
- ‘Filters’ specifies flags used to filter out the variants containing them.
- ‘Samples’ provides options for the variants track containing variants of multiple individuals. Each selected individual is displayed in one track. If no individual options are selected, all individuals’ variants are displayed in one track.

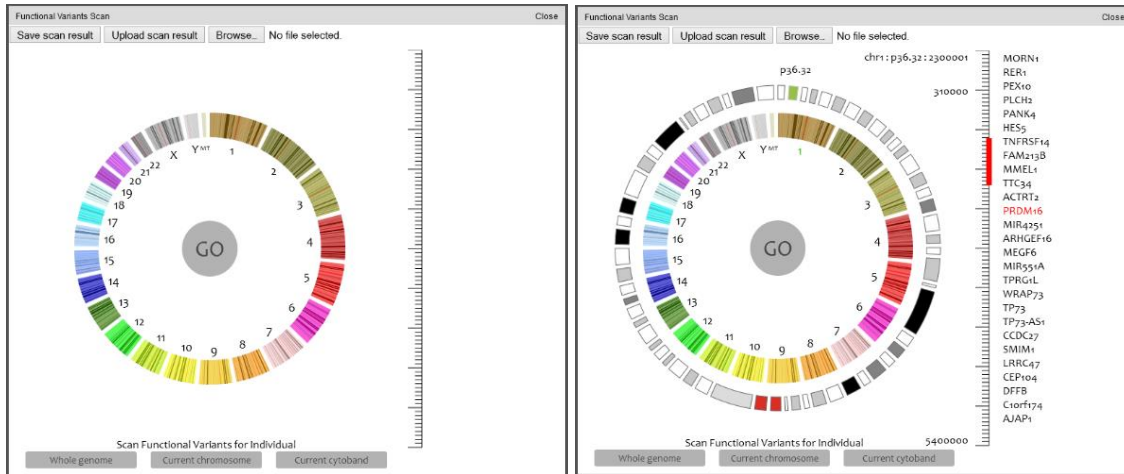
Functional variants scan

Clicking on the ‘Scan’ button on the right of the chromosome ideogram, then the functional variants scan window pops up.

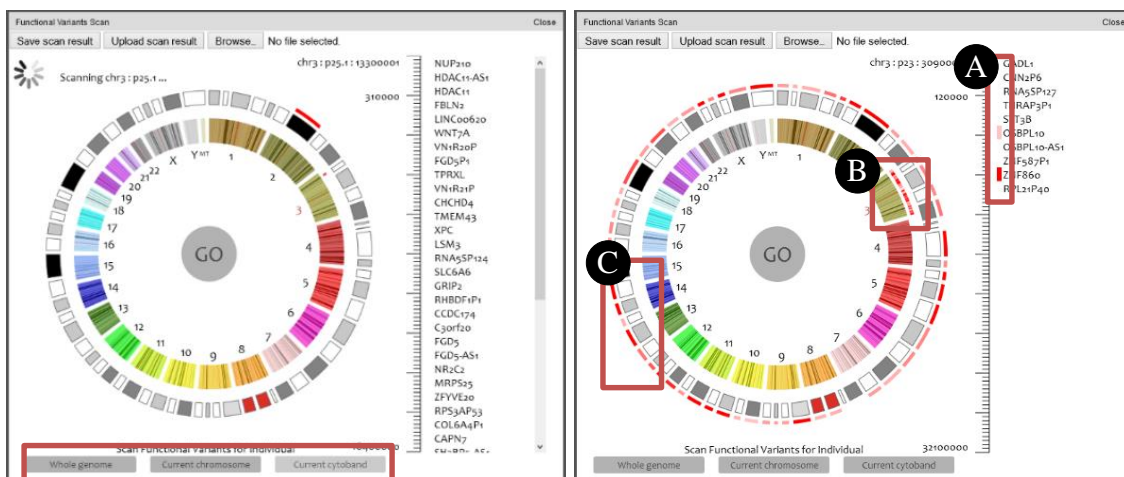


A ‘genome map’ is displayed in the pop-up window. The map consists of ‘chromosomes ring’ shown in the inner circle, ‘cytobands ring’ shown in the outer circle, and the gene list displayed in the right area. The inner circle ‘chromosomes ring’ displays all human chromosomes. Moving or clicking the cursor to select a chromosome, the ‘cytobands ring’ of the selected chromosome is displayed in the outer circle. Moving the cursor over a cytoband on the cytobands ring, the id of the cytoband appears nearby. After clicking on the cytoband, the genes located in the cytoband are listed on the right of the window. The vertical ruler on the left of the gene list indicates the length of the selected cytoband. The chromosome name, cytoband id, and starting coordinate of the cytoband are displayed close to the first major tick of the ruler. And the ending coordinate is displayed close to the last major tick of the ruler. The length of the genomic range corresponding to the

ruler's intervals of two major ticks is shown close to the second major tick. Moving the cursor over the gene list, a red line that marks the gene's position and length appears on the ruler. Users can click a gene to view it in the browser view.

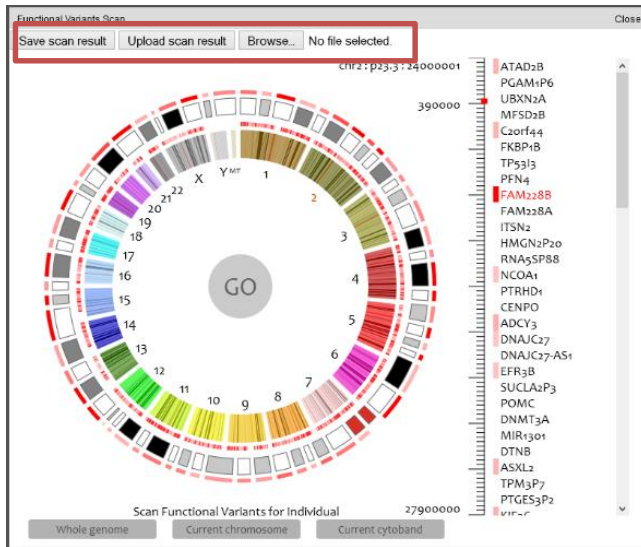


Three ranges, including a cytoband, a chromosome and the whole genome, are provided for scanning functional variants at the bottom of the window. The scanning time depends on the length of the region, the number of the genes and individual genome variants in the region. The scanning result is displayed during the scanning process on time. Based on the scanning result, the disorder risk levels of genes and cytobands are colored nearby as shown in area A, B and C. Users can stop the scanning process at any time by closing the functional variants scan window and the scanning result is recorded in the server. When reopening the closed window, the scanning result is displayed again.



The 'save scan result' button on the top of the functional variants scan window is used to download and save the result file to the user's local disk. The file can be uploaded to the PGB by clicking 'upload scan result' button later on. Users can also upload revised result file to display the customized heatmap of genome disorder risk.

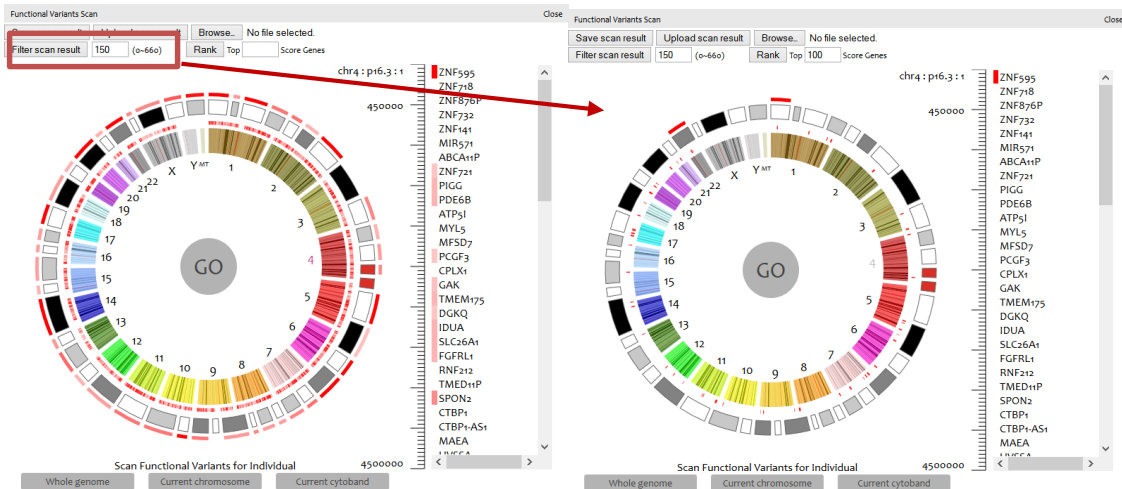
The file has five columns: chromosome, starting coordinate, ending coordinate, cytoband id/gene symbol, and score. The first four columns of each row indicate the range and the name of a cytoband region or a gene region, and is not allowed to be changed by users. Users can change the values of the last score column. The values, which range from 0 to 100, determine the heatmap color of the region indicated by the first four columns. The scanning result for the default individual genome data (*i.e.* NA12716 genome data from the 1000 Genomes Project), has been initialized in the PGB.



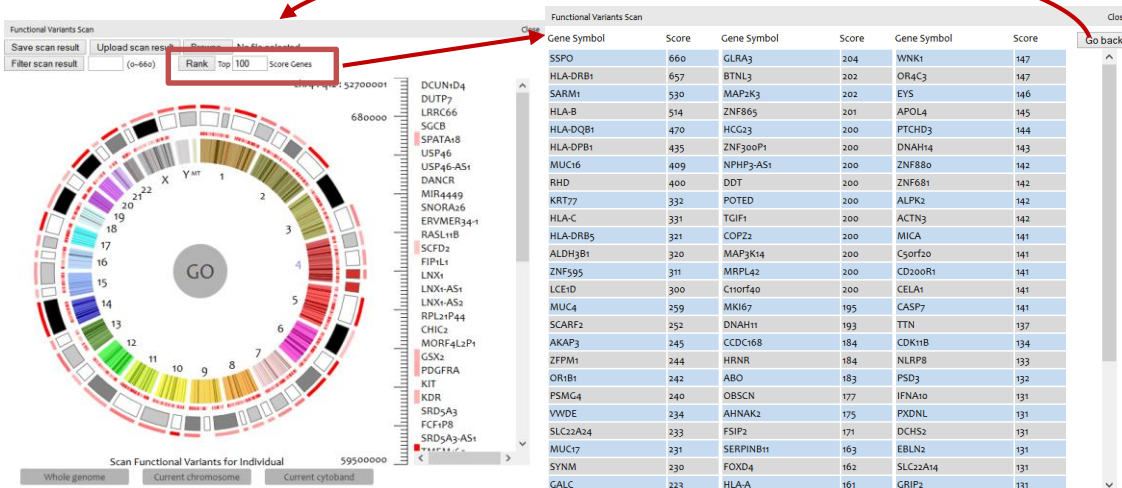
Gene filtering and ranking

The PGB provides gene ranking and filtering functions to facilitate users to investigate the individual variants.

- a) A ranking function is provided in the Functional Variants Scan window of PGB, to rank functional effects of genes by scores. The ranking result is displayed in descent order in a pop up window based on the scanned result. The number of genes to be ranked and displayed can be specified by users.
- b) A filter function is also provided in the Functional Variant Scan window of PGB. Given a user specified threshold, only genes and cytobands with scores over the threshold are labelled in red in the circular heatmap. Users can easily focus on interested genes by filtering out genes with lower functional effects scores.



(a)



(b)

Track management

The tracks in the reference genome panel can be controlled selectively by checking the checkboxes in the track management panel. Each drop-down menu in the panel groups built-in genomic annotation or individual genome data. Users can collapse/expand all grouped contents by clicking the ‘collapse all’ or ‘expand all’ buttons. The users provided tracks are grouped in ‘custom track’ drop-down menu (Figure S5C).

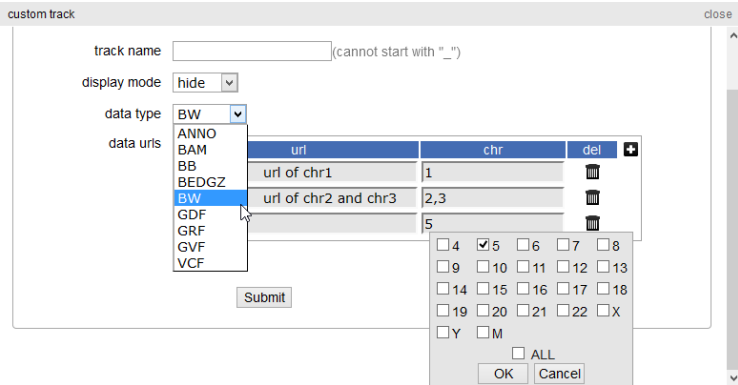
To add/remove/change the tracks in the individual genome panel, users can click the ‘setting’ button (Figure S5B) to pop up the personal tracks setting window. The window contains four drop-down menus: Individual variant scoring methods, Individual

Functional Elements, Individual Genes and Individual Phenotypes. The top-3 menus provide radio buttons and single option menus, and the bottom one supports checkboxes and the multiple option menu. The selected annotations are displayed in the individual genome panel. To make sure to select the items in the phenotype drop-down menu, the options in the top three drop-down menus are necessary. The ‘clear’ button on the bottom of a drop-down menu allows users to clear all selected options in that menu.

Users can pop up the ‘Add Custom Tracks’ window (Figure S5A) by clicking on the ‘Add Custom Tracks’ button on the menu bar. There are three tabs in the custom track window: add one track, add multiple tracks, and remove custom tracks.

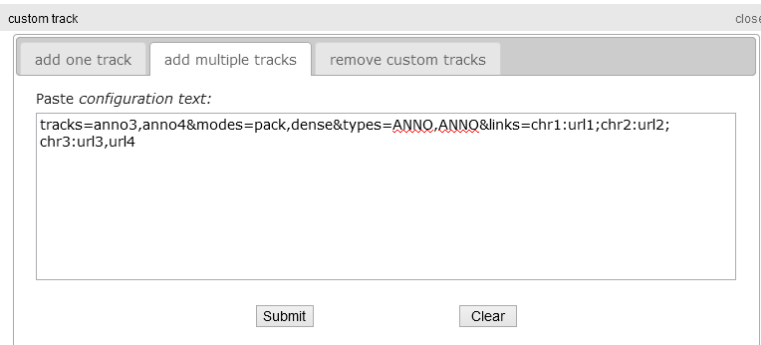
In order to upload data for a custom track, the following four parameters are required.

- Track name: track name is not allowed to start with underscore ‘_’, or to be same as any existing track name. After adding the custom tracks, the track names are grouped in the track management panel.
- Display mode: there are three option for the parameter: hide, dense and pack. After uploading the custom track data, the custom track is hided in hide mode, and the custom track is displayed in dense/pack mode in the reference genome panel.
- Data type: The PGB supports 9 types of data and annotation files, including BAM, BB (BigBed), BW (BigWig), ANNO, GDF, GRF, GVF, and VCF (**Table 1** in Article). ANNO, GDF, GRF, GVF and VCF files should be compressed and indexed by Tabix tool.
- URLs: Users can specify URLs of chromosome data files. Users can add/remove chromosomes and URL pairs for a track by clicking on the ‘plus’/‘trash bin’ button.



Users can upload multiple tracks data by inputting the configuration text in the second tab. The configuration text contains four parameters: ‘tracks’, ‘modes’, ‘types’ and ‘links’, which correspond to four parameters in the first tab. In the text:

- ‘&’ is the separator of the four parameters,
- ‘=’ connects keys and values of each parameter,
- ‘,’ is the separator of parameter values for different tracks. The track values should be inputted in the same order of the parameters, and
- ‘links’ parameter can contain multiple chromosomes-URL pairs, where ‘;’ is used to separate different chromosome-URL pairs, ‘:’ connects chromosomes and URL of each pair.



In the remove custom track tab, users can select the checkboxes and click ‘remove selected’ button to remove selected custom tracks.

